Interprovincial differences in the rates of minor crimes of violence and related disorders in New Zealand 1853-1930: part 2

Haslett, Stephen J.; Fairburn, Miles

Historical Social Research, Vol. 16 — 1991 — No. 3, 40-68

Interprovincial Differences in the Rates of Minor Crimes of Violence and Related Disorders in New Zealand 1853 - 1930:

Part II

*Stephen Haslett, Miles Fairburn**

Abstract: The problem of comparing historical data with common variables from two or more distinct locations remains an open question in historical studies. The issues of formulating suitable historical models and comparing them, using appropriate statistical models and techniques, are the topics of this paper. These matters are first discussed in general and a number of possible techniques outlined in concept. The advantages and disadvantages of each are summarised. The question of distinguishing between differences of structure and differences of degree, in the presense of measurement error, is then considered in greater detail with reference to factor analysis and the New Zealand provincial data base, 1853 - 1930.

1. Introduction

In an earlier contribution to HSR (Vol. 1 Haslett and Fairburn, 1990) we asked whether the ecological structure of minor violent crime in New Zealand from 1853 to 1930 was subject to a large amount of regional variation. We found that in each of New Zealand's nine regions the structure differed not just in degree but in type (or kind) from that for all the nine regions taken as a whole and that each region had a different type of structure. We also found, however, that the structural variations were not fundamental in type, i.e. each regional structure was a sub-type of the

* Address all communications to Stephen Haslett, Institute of Statistics and Operations Research, Victoria University of Wellington, Wellington, New Zealand; or to Miles Fairburn, History Department, Victoria University of Wellington, Wellington, New Zealand.

global model. On this basis, we speculated that the same model may have been universal to other frontier societies. This model we called 'atomism': a deficiency in the means of association and informal regulation.

The purpose of the following article is to spell out in detail the technical and statistical underpinnings of the earlier article. We hope the discussion might prove useful since it addresses a wider problem that is often struck in quantitative social historical research - the problem of how to decide whether the same model fits two or more sets of data having common variables.

Part II of this article is intended to be rather more a methodological note, than further exposition of the historical material on interprovincial differences in crime rates and related disorders in New Zealand, 1863-1930. The intention is to provide sufficient statistical detail that the two parts of the article, taken together, might provide a blueprint for comparative historical studies involving measurement error, where the distinction between difference in structure and difference of degree needs to be made.

For this reason tables and graphs referred to, but not contained in Part I, have the same titles and numbering system in Part II. Part II thus forms the complement to Part I, not only in that it contains the graphs and figures referenced but not given in the first part, but also in that it provides the statistical justification for the analytical methods used to analyse the multivariate time series data for each of the nine New Zealand provinces and to compare their underlying structure.

2. Statistical Techniques - General Notes

Some very general background notes on statistical techniques appropriate for analysis of historical data in the form of time series seem appropriate here. Particular consideration is given to multiple regression, multivariate time series techniques, principal components analysis, measurement error models, and factor analysis. The discussion is developed without the use of mathematical formulae.

2.1 Multiple Linear Regression

Multiple linear regression is a useful technique for using linear combinations of a set of independent variables to find, separately, a best estimate for each of a set of dependent variables. See, for example, Dunn and Clark (1974), or Seber (1977). The technique is essentially a predictive one; it seeks to minimise the sums of the squares of the differences between the observed and estimated dependent variable. The regression coefficient for a particular independent variable (and a given dependent variable) measures the predicted effect on the dependent variable of a unit change in that particular independent variable, with all other independent variables being held constant.

A number of cautions about the use of multiple regression are both warranted and necessary.

- Regression techniques assume that each independent variable has only a linear effect on the dependent variable; squares of a variable and the product of two variables can, for example, be determined and included specifically as independent variables which themselves have linear effects, but this often creates problems of multicollinearity of independent variables, variable selection and interpretation of regression coefficients.
- Model errors, or the differences between observed and expected values of the dependent variable, are assumed not to be correlated, for example over time. Caution is necessary then when using regression techniques on time series data, especially where rates of change or growth rates are included in the independent variables. Correlation of model errors over time (i.e. autocorrelation) makes the usual least squares estimates of regression coefficients inefficient (i.e. such estimates are no longer best linear unbiased estimates).
- Interpretation of regression models in terms of causes is fraught with difficulty, because the technique makes no attempt to determine the interdependences between variables except in a predictive sense. Multiple regression coefficients are consequently often poor indicators of the relative importance of independent variables in some underlying structural model.
- All independent variables are assumed to be observed without error. If there is measurement error in these variables, estimates of regression coefficients will be biased.

In summary, multiple (linear) regression is a useful and appropriate predictive technique where data are not autocorrelated, independent variables are observed without error, and the dependent variable can be modelled as a linear combination of the independent variables.

2.2 Multivariate Time Series Techniques

Multivariate time series analysis is a very broad field. See, for example, Hannan (1970), Koopmans(1974), Box and Jenkins (1976), Priestley (1981, 1988). While techniques include fitting non-linear and non-stationary time series (e.g. Priestley, 1988), much of the necessarily elaborate theory has been developed for the simpler case of linear stationary time series (Koopmans (1974, pp. 38, 81)). (A time series is, for example, (weakly) stationary if it has a constant mean and variance.) Under fairly general conditions, namely the continuity of its spectrum, every weakly stationary process has an infinite order moving average representation. This leads rather naturally, via rational spectra, into ARMA (autoregressive-moving average) representations for time series, in which a process in time is viewed as a linear combination of its own past values (hence auto-regressive) and the present and past values of a white noise, or self-uncorrected, process (hence moving average). When exogenous (or independent) variables are also available ARMAX models are appropriate. In ARMAX models the dependent variable is regressed on a number of independent variables as in multiple regression, but the model error rather than being independent and identically distributed is an ARMA process (see, for example, Judge et al. (1980, Chapters 5, 6)). ARMAX models extend multiple regression models in allowing the error from the regression to be autocorrelated, but otherwise the cautions mentioned above for regression models apply. Even in its more general context of non-stationary and/or non-linear time series, the techniques are essentially predictive, and ARMAX models assume that exogenous variables are observed without error.

2.3 Principal Components Analysis

Principal components analysis (see, for example, Morrison (1976), Seber (1984), Flury (1988)) differs from regression and multivariate time series techniques in that no distinction is made between dependent and independent variables. If variables are rescaled to have unit variances, the analysis becomes a principal components analysis of the correlation matrix. If a multidimensional scatterplot of the (possibly rescaled) data points were made (with the same number of dimensions as there are variables) then the first principal component would correspond to the direction in which the data has maximum spread within this multidimensional space, the second principal component to the direction of next greatest spread (conditional on that direction being orthogonal (i.e. at right angles) to the first principal component), etc. There are in general as many eigenvectors (or principal components) as there are variables, and the extent of spread in the direction of each eigenvector is measured by the corresponding eigenvalue. In general, the covariance and correlation matrices will have different eigenvalues and different eigenvectors.

Principal components involves a decomposition of the covariance or correlation matrix. Once this decomposition is achieved a principal component score corresponding to each eigenvalue can be calculated for each observation. These scores measure the position of the data point on each particular principal component axis, i.e. then the principal components are used as a reference system. Finding eigenvectors and eigenvalues corresponds to a rotation of the original orthogonal axis system to a new system, also orthogonal, in which scores on one component (or axis) are uncorrected with scores on any other.

When the original variables are highly inter-correlated, the first few principal components often explain much of the variability or spread in the data. If the data are variables observed over time, a plot of the principal component scores over time may reveal much about the underlying structure of the data.

Further, if the time series data are stationary in each component series, principal components can be found at various frequencies in the spectrum of the multivariate time series (Brillinger (1975), Chapter 9, especially p.353) and this may reveal component trends (at low frequencies) or common short-term patterns (at high frequencies). An analogous procedure is simply to include past (i.e. lagged) values of the component series in the vector for which the covariance (or correlation) matrix is calculated. Adding lagged variables is of little benefit however when lag correlations are close to zero or close to one.

Although often illuminating, use of principal components analysis is not without its own possible complications:

- Principal components analysis assumes the model is linear in the variables.
- The scaling of variables is critical; unless variables are on a somewhat similar scale the variables with the greatest spread will predominate in the principal components. Scaling of variables can be somewhat arbitrary when variables are on different scales; Morrison (1976, p.268) suggests that a principal component analysis of the correlation matrix should be considered in such circumstances.
- If the variables are observed with measurement error, the eigenvalues and eigenvectors of the covariance (or correlation) matrix of the measured variables are not those of the variables measured without error.
- Interpretation of principal components, particularly by giving them names, does not remove the fact that they are chosen simply as linear combinations of variables having maximum spread (possibly subject to orthogonality constraints imposed by other eigenvectors).

Principal components analysis is nevertheless a useful data reduction technique. Where the data being analysed can be interpreted as belonging to groups each of which gives rise to a covariance matrix, the existence of common principal components between groups may be studied using the methods outlined in Flury (1988).

2.4 Measurement Error Models

Of the three statistical methods discussed previously, multiple regression, multivariate time series techniques, and principal components, it is only certain (multivariate) time series techniques that allow explicitly for possible measurement error. Indeed, even for time series modelling and ARMAX models the exogenous (independent) variables must be observed without error. An extensive theory detailed in Fuller (1987) has been developed for regression in the presence of measurement error but such techniques require knowledge of the measurement error variances and little suitable computer software is at present available. If measurement error variances are known, and measurement error is assumed uncorrected for different variables (and over time) then principal components analysis is made possible by subtracting the measurement error variances off the appropriate diagonal elements of the covariance matrix before beginning the principal component analysis. When the measurement error is known and structured in this way, the off-diagonal elements of the covariance matrix are unaffected by the measurement error.

2.5 Factor Analysis

Factor analysis has been a vexed subject, at least in the statistical literature (see, for example, Seber). The principal difficulties centre around existence of a factor solution for a given covariance or correlation matrix, criteria for specifying the number of common factors, convergence of the maximum likelihood algorithm for estimating factors given multivariate normal data, and the naming of factors. Two distinct advantages of factor analysis are that it is unaffected by uncorrected measurement error, and that rescaling of variables affects the model only through a rescaling of the factor loadings.

Factor analysis is, like principal component analysis, a decomposition of the covariance or correlation matrix but the two techniques are in no way equivalent. Factor analysis attempts to find a number of common factors (or latent variables) such that given these common factors, the original variables in the analysis are conditionally independent. These factors may be orthogonal or oblique, i.e. factor analysis is a technique for analysing a covariance or correlation matrix to see whether it can be decomposed into a positive definite matrix of rank equal to the number of factors, and a diagonal matrix of specific variances; the second matrix in the decomposition contains specific variances as diagonal elements and has zero off-diagonal elements due to the conditional independence of variables given the factors. There is no guarantee that such a decomposition exists for a given covariance (or correlation) matrix.

Factor analysis has been popularised, at least in the social science literature, by the availability of the LISREL package (Joreskog and Sorbom, 1986) which allows linear structural relationships to be fitted to multivariate data. Kiiveri and Speed (1982) have cautioned that such a model fitting routine is best used to fit a model suggested on a priori grounds rather than to initiate an extensive model search.

Where time series data is being analysed, and the correlation of lagged values (i.e. the autocorrelation) of particular variables is not near zero or near one, lagged variables may be included in the covariance or correlation matrix used in the factor analysis. Brillinger (1975, p.353) briefly discusses factor analysis of time series data in the frequency domain. Differenced variables or measures of rates of change of variables can also be included in the covariance (or correlation) matrix analysed. (See, for example, Fairburn and Haslett, 1986.) Morrison (1976) notes that factor analysis (like multiple linear regression, linear time series techniques and principal components analysis) involves a search for a model which is linear in the variables under study, and that lack of linearity in the underlying model is as proper a conclusion if a factor model does not fit sample data as is the conclusion that a factor model does not exist.

2.6 Summary

By way of summary then, each of these statistical techniques (multiple linear regression, (multivariate) time series analysis, principal components, measurement error models, and factor analysis) is appropriate in certain circumstances. Nearly all these techniques assume that the model is linear in the variables under study. Multiple regression and time series techniques are essentially predictive; principal components and factor analysis allow fitting of models with possibly greater explanatory power. Generally speaking, only measurement error models and factor analysis can be used where variables are measured with error.

3. Application to the New Zealand Data

The variables analysed in this study consist of the rate variables listed in Appendix 1. An outline of the particular methods of analysis listed here, including confirmatory factor analysis, is given in Section III of Part I. When supplemented with the additional material in the Appendices, these together detail the methods of analysis used for this study, and, we hope, provide a blueprint for any future studies where data involving the same variables collected from different regions over time are to be compared to make a distinction between differences of degree and differences of structure.

References

Box, G E P and Jenkins, G M (1976). Time Series Analysis : Forecasting and Control. Holden Day.

Brillinger, D R (1975). Time Series : Data Analysis and Theory. Holt, Rinehart and Winston Inc., New York.

Dunn, O J and V A Clark (1974). Applied Statistics : Analysis of Variance and Regression. John Wiley & Sons, New York.

Fairburn, M and Haslett, S J (1986) Violent Crime in Old and New Societies - A Case Study Based on New Zealand 1853-1940. Journal of Social History, vol 20, no. 1, pp 89-126.

Flury, B(1988). Common Principal Components and Related Multivariate Models. John Wiley & Sons, New York.

Fuller, W A (1987). Measurement Error Models. John Wiley & Sons, New York.

Hannan, E J (1970). Multiple Time Series. John Wiley & Sons, New York.

Haslett, S J and Fairburn, M (1990) Interprovincial Differences in the Rates of Minor Crimes of Violence and Related Disorders in New Zealand, 1853-1930. Historical Social Research, vol. 15, no. 4, pp. 140-183.

Joreskog, K G and D Sorbom (1986). LISREL VI User's Guide, 4th edition. Scientific Software Inc., Indiana.

Judge, G G, Griffiths, W E, Hill, R C and T-C Lee (1980). The Theory and Practice of Econometrics. John Wiley & Sons, New York.

Kiiveri, H and T P Speed (1982). Structural Analysis of Multivariate Data : A Review. In Sociological Methodology, 1982, pp.208-289. Edited by Samuel Leinhardt. Jossey-Bass.

Koopmans, L H (1974). The Spectral Analysis of Time Series. Academic Press, New York.

Morrison, D F (1976). Multivariate Statistical Methods, 2nd edition. McGraw-Hill, London.

Priestley, M B(1981). Spectral Analysis and Time Series. Academic Press, New York.

Priestley, M P (1988). Non-linear and Non-stationary Time Series Analysis. Academic Press, New York.

Seber, G A F (1977). Linear Regression Analysis. John Wiley & Sons, New York.

Seber, G A F (1984). Multivariate Observations. John Wiley & Sons, New York.

APPENDIX 1

Description of Variables Analysed

Percentage of Total Population who are Overseas-bom Europeans
Percentage of Total Population who are Irish-born
Percentage of Total Population who are European Adult Males
Percentage of Total Population who are European Adult Females
Percentage of Total Population who are European and Urban
Percentage of Total Population who are Young European Males (aged 21-40 years)
European Dwellings 1-2 Rooms, per capita
European Total Dwellings, per capita
Spirits, Imperial Gallons Consumed, per capita
Beer, Imperial Gallons Consumed, per capita
Ratio of European Adult Males to Adult Females
European Dwellings, 1-2 Rooms, as a Ratio of Total Dwellings
Imports, pound per capita
Exports, pound per capita
Manufacturing Horse Power per capita
Percentage of Total Population who are Manufacturing Employees
Police Manpower, per 100,000 of Total Population
Immigration rate, per 100.000 of Total Population
Emigration rate, per 100,000 of Total Population
Drunkenness rate, summary convictions, per 100,000 of Total Populalation
Civil Suits tried and disposed of in the Magistrates' Courts, rate per 100,000 of Total Population
Violence rate, per 100,000 of Total Population

Notes: Details of the analysis period, population bases, and data sources can be found in Appendix 1 of Fairbum and Haslett (1986).

APPENDIX 2A

The Effect of Measurement

Factor analysis involves the decomposition of a p-dimensional covariance or correlation matrix $\mathbf{\Sigma}$ into the product with its transpose of a factor loading matrix $\mathbf{\Lambda}$ of rank $k < p$, and a diagonal matrix of specific variances

$$\mathbf{\Sigma} = \mathbf{\Lambda}\mathbf{\Lambda}^T + \mathbf{\Psi} \tag{A1}$$

where $\mathbf{\Lambda}$ is the loading matrix of rank k

$\mathbf{\Lambda}^T$ is $\mathbf{\Lambda}$ with the rows of $\mathbf{\Lambda}$ as columns (and the columns as rows)

$\mathbf{\Psi}$ is a diagonal matrix containing specific variances.

From this equation it is clear, since $\mathbf{\Psi}$ is diagonal, that the p variables for which $\mathbf{\Sigma}$ is the covariance (correlation) matrix are conditionally independent given the factors.

We denote these p variables by the vector x, and consider each realisation of x to be the true x, namely x° plus a random error e, i.e.

$$\mathbf{x} = \mathbf{x}^0 + \mathbf{e} \tag{A2}$$

If the components of the error e are taken to be uncorrelated with themselves and with x° then the covariance (correlation) matrix of e is itself diagonal.

Further the covariance (correlation) matrix of x can be written

$$\begin{aligned}\mathbf{\Sigma_x} &= \mathrm{E}[(\mathbf{x} - \boldsymbol{\mu}_\mathbf{x})(\mathbf{x} - \boldsymbol{\mu}_\mathbf{x})^T] \\ &= \mathrm{E}[(\mathbf{x}^0 + \mathbf{e} - \boldsymbol{\mu}_\mathbf{x})(\mathbf{x}^0 + \mathbf{e} - \boldsymbol{\mu}_\mathbf{x})^T] \\ &= \mathrm{E}[(\mathbf{x}^0 + \boldsymbol{\mu}_{\mathbf{x}^o} - \boldsymbol{\mu}_{\mathbf{x}^o})^T] + \mathrm{E}(\mathbf{e}\,\mathbf{e}^T) \\ &= \mathbf{\Sigma}_{\mathbf{x}^o} + \mathbf{\Delta}, \text{ say}\end{aligned} \tag{A3}$$

where $\mathbf{\Delta}$ is a diagonal matrix. Here E denotes expected value. Note that the cross product term $\mathrm{E}[\mathbf{x}^o - \boldsymbol{\mu}_{\mathbf{x}^o})$ has expectation zero and hence vanishes, since e and x° are assumed uncorrelated; note also that the population mean of x, namely $\boldsymbol{\mu}_\mathbf{x}$, equals the population mean of x°, namely $\boldsymbol{\mu}_{\mathbf{x}^o}$, since the mean of e , E(e) equals zero.

The importance of the above equation is that, providing the e are not correlated with

themselves or with the true values x°, the off-diagonal elements of the covariance matrix of x are the same as the corresponding off-diagonal elements of the covariance matrix of x°. A factor analysis of Σ_x will thus yield the same factors as a factor analysis of Σ_{x° since Δ will be absorbed into the diagonal matrix specific variances, i.e.

$$\Sigma_x = \Lambda_{x^\circ}\Lambda^T_{x^\circ} + \Psi_x \qquad \text{(A4)}$$

where

$$\Psi_x = \Psi_{x^\circ} + \Delta \qquad \text{(A5)}$$

and where Λ_{x° is the factor loading matrix for the true values x°.

In these circumstances then, the factors are unaffected by measurement error.

APPENDIX 2B
On the Relationship Between the Global Analysis and Separate Factor Analyses for Each Province

In Fairburn and Haslctt (1986) a factor analysis model was fitted to the combined data from all nine provinces. In the present paper we consider separate factor analysis models for each province. This appendix explores the mathematical relationship between the two analyses.

Given a vector of variables x (such as those in Appendix 1) we may write

$$\mathbf{x} - \boldsymbol{\mu}_\mathbf{x} = \boldsymbol{\Lambda}\boldsymbol{\xi} + \boldsymbol{\delta} \tag{B1}$$

where $\boldsymbol{\mu}_\mathbf{x}$ is the (population) mean of x. (In fact in most factor analyses x is implicitly "mean corrected" so that it is $(\mathbf{x} - \boldsymbol{\mu}_\mathbf{x})$ that is analysed rather than x; if the correlation matrix is factor analyses, $(\mathbf{x} - \boldsymbol{\mu}_\mathbf{x})$ is also standardised by dividing by the appropriate standard deviations.)

By multiplying equation (Bl) by its transpose and taking expectations

$$\begin{aligned}
\boldsymbol{\Sigma}_\mathbf{x} &= \mathrm{E}[(\mathbf{x} - \boldsymbol{\mu}_\mathbf{x})(\mathbf{x} - \boldsymbol{\mu}_\mathbf{x})^\mathrm{T}] \\
&= \mathrm{E}[(\boldsymbol{\Lambda}\,\boldsymbol{\xi} + \boldsymbol{\delta})(\boldsymbol{\Lambda}\,\boldsymbol{\xi} + \boldsymbol{\delta})^\mathrm{T}] \\
&= \boldsymbol{\Lambda}\,\mathrm{E}(\boldsymbol{\xi}\,\boldsymbol{\xi}^\mathrm{T})\boldsymbol{\Lambda}^\mathrm{T} + \mathrm{E}(\boldsymbol{\delta}\,\boldsymbol{\delta}^\mathrm{T}) \\
&= \boldsymbol{\Lambda}\boldsymbol{\Lambda}^\mathrm{T} + \boldsymbol{\Psi}
\end{aligned} \tag{B2}$$

since $\mathrm{E}(\boldsymbol{\xi}\,\boldsymbol{\xi}^\mathrm{T}) = \mathbf{I}$, the identity matrix, and
$\mathrm{E}(\boldsymbol{\delta}\,\boldsymbol{\delta}^\mathrm{T}) = \boldsymbol{\Psi} = \boldsymbol{\Psi}_\mathbf{x}$ of Appendix 2A.

Noting that $\boldsymbol{\Lambda}$ here is $\boldsymbol{\Lambda}_\mathbf{x} = \boldsymbol{\Lambda}_{\mathbf{x}0}$ of Appendix 2A, equations (Bl) and (B2) provide the connection between the factor scores $\boldsymbol{\xi}$ and the decomposition of the covariance (correlation) matrix.

Some specifications of the factor analytic model decompose $\boldsymbol{\Sigma}_\mathbf{x}$ as

$$\boldsymbol{\Sigma}_\mathbf{x} = \boldsymbol{\Lambda}_1\boldsymbol{\Phi}\boldsymbol{\Lambda}_1^\mathrm{T} + \boldsymbol{\Psi} \tag{B3}$$

(see, for example, Jõreskog and Sorbom(1986)). Here $\boldsymbol{\Phi}$ - $\mathrm{E}(\boldsymbol{\xi}\,\boldsymbol{\xi}^\mathrm{T})$, so that if $\boldsymbol{\Phi} = \mathbf{I}$, the identity matrix, we get equation (B2), and if $\boldsymbol{\Phi}$ has non-zero off-diagonal elements (with the columns of $\boldsymbol{\Lambda}_1$ orthogonal) we get an oblique factor analysis. By setting $\boldsymbol{\Lambda} = \boldsymbol{\Lambda}_1\boldsymbol{\Phi}^{1/2}$ where $\boldsymbol{\Phi}^{1/2}\boldsymbol{\Phi}^{1/2} = \boldsymbol{\Phi}$, equations (B2) and (B3) are seen to be equivalent.

The connection between a factor analysis of the correlation matrix and that for the covariance matrix is provided by

$$\begin{aligned} R_x &= D^{-1}\Sigma_x D^{-1} \\ &= D^{-1}(\Lambda\Lambda^T + \Psi) D^{-1} \\ &= (D^{-1}\Lambda)(D^{-1}\Lambda)^T + D^{-1}\Psi D^{-1} \\ &= \Lambda_{R_x}\Lambda_{R_x}^T + \Psi_{R_x} \end{aligned} \qquad \text{(B4)}$$

where D is a diagonal matrix containing standard deviations as diagonal elements, and D^{-1} is its inverse. Thus the factor analytic decomposition of a correlation matrix yields the same factors as the decomposition of the corresponding covariance matrix, when the factor loading matrix of the covariance matrix are scaled by dividing by the appropriate standard deviations. (See Everitt (1984), equation (2.6), p. 16 for further details.)

Fairburn and Haslett (1986) contained a factor analytic model fitted to the correlation matrix of a subset of the variables in Appendix 1; the factor model for decomposition of the covariance matrix of the same variables is related to the factor model for the correlation matrix by equation (B4). Only the first factor was considered important in that analysis.

The model of Fairburn and Haslett (1986) is given by equation (B1) with Λ containing a single factor (so that A is itself a column vector). The nine provinces arc known to have different sample means (and by inference different population means) for the same variables, so that $\mu_x^{(g)} \neq \mu_x$ for different provinces g • 1,2,...,9 where $\mu_x^{(g)}$ is the (population) mean of x in province g.

Now a separate factor analysis applied to the covariance matrix in each province fits

$$x^{(g)} - \mu_x^{(g)} = \Lambda^{(g)}\xi^{(g)} + \delta^{(g)} \qquad \text{(B5)}$$

We seek the relationship between equations (B1) and (B5) when equation (B1) applies to the combined provincial (i.e. the global) data.

From equation (B1)

$$x^{(g)} - \mu_{x^{(g)}} = (\mu_x - \mu_{x^{(g)}}) + \Lambda\,\xi^{(g)} + \delta^{(g)} \qquad \text{(B6)}$$

where the superscript, g, is used to denote that the random variables, or parameters, are those relating to province g. (Note that $\xi^{(g)}$ and $\delta^{(g)}$ are simply those realisations of ξ and δ of equation (B1) for province g.)

Under the condition that the error process in equation (B6) has zero mean not only overall but also within each province, we have

$$E(\delta^{(g)} \mid g) = E(\delta^{(g)}) = 0$$

i.e. that the population mean of the error process generating the specific variances be zero for each province considered separately.

By way of simplification of notation, let

$$\mu_{\xi^{(g)}} = E(\xi^{(g)} \mid g) \qquad \text{(B7)}$$

$$\mu_{\xi^{(g)}} = E(\xi^{(g)} \mid g) \qquad \text{(B7)}$$

and

$$\mu_x^{(g)} - \mu_x = \Lambda \, \mu_{\xi^{(g)}} \qquad \text{(B9)}$$

$$\Sigma_{\xi^{(g)}} = \text{variance}(\xi^{(g)} \mid g) \qquad \text{(B8)}$$

For simplicity we now consider the case in which the factors taken globally are orthogonal, and the factors in each province are also orthogonal. $\Sigma_{\xi^{(g)}}$ and Σ_{ξ} are then diagonal (where

$$\Sigma_{\xi} = E(\xi \, \xi^{T}) = I \qquad \text{(B10)}$$

Then taking expectations of equation (B6) for each province, g, separately, yields

is the unconditional covariance matrix for ξ). If we now standardise the random variables $\xi^{(g)}$ to have uncorrelated components with zero mean and variance one, using

$$\xi_s^{(g)} = \Sigma_{\xi^{(g)}}^{-1/2} (\xi^{(g)} - \mu_{\xi^{(g)}}) \qquad \text{(B11)}$$

we have, on rea

$$x^{(g)} - \mu_{x^{(g)}} = (\mu_x - \mu_{x^{(g)}}) + \Lambda \, \mu_{\xi^{(g)}} + \Lambda \, \Sigma_{\xi^{(g)}}^{-1/2} \, \xi_s^{(g)} + \delta^{(g)} \qquad \text{(B12)}$$

$$x^{(g)} - \mu_{x^{(g)}} = (\mu_x - \mu_{x^{(g)}}) + \Lambda \, \mu_{\xi^{(g)}} + \Lambda \, \Sigma_{\xi^{(g)}}^{-1/2} \, \xi_s^{(g)} + \delta^{(g)} \qquad \text{(B12)}$$

and hence from equation (B9) that

$$x^{(g)} - \mu_{x^{(g)}} = (\Lambda \Sigma_{\xi^{(g)}}^{-1/2}) \xi_s^{(g)} + \delta^{(g)} \qquad \text{(B13)}$$

Equation (B13) is then the factor analytic model for each province, g, corresponding to the global factor analytic model of equation (B1) under the assumption that the mean of the error process in equation (B13) has zero mean for each province, g. Equation (B13) has an elegant interpretation: it is a factor analytic model for which the factor loading matrix for each province, g, is simply a rescaling of the factor loading matrix from the global analysis. Each column of the global factor loading matrix is scaled for province g, by the corresponding diagonal element of the square root of the covariance matrix $\Sigma_{\xi(g)}$ (i.e. the jth column of Λ is scaled up by the standard deviation of $\xi^{(g)}$).

References

Everitt, B S (1984). An Introduction to Latent Variable Models, Chapman and Hall, London.

Jöreskog, K G and Sörbom, D (1986). Lisrel VI: User's Guide, 4th edition, Scientific Software Indiana.

APPENDIX 3

Additional Tables

TABLE A-I

Measures of goodness of fit for constrained and unconstrained factor analyses by province

	Unconstrained			Constrained		
	Adjusted goodness of fit index	Root mean square residual	% of normalised residuals>3 *	Adjusted goodness of fit index	Root mean square residual	% of normalised residuals>3 *
Auckland	0.918	0.040	3.5	0.486	0.110	62.0
Hawkes Bay	0.978	0.029	1.0	0.790	0.097	41.0
Taranaki	0.977	0.015	5.5	0.592	0.069	48.5
Wellington	0.933	0.070	3.5	0.414	0.228	75.0
Marlborough	0.971	0.013	5.5	0.581	0.054	55.0
Nelson	0.993	0.016	1.0	0.664	0.118	69.5
Canterbury	0.946	0.041	6.0	0.743	0.098	53.5
Westland	0.947	0.105	5.0	0.912	0.148	17.0
Otago	0.921	0.055	11.5	0.815	0.093	27.5

Notes: For constrained fits the Q plot had an average slope of approximately 0.3, indicating poor fits, with very little variation of slope from province to province.

For unconstrained fits the Q plot had an average slope of approximately one, indicating moderately good fits, with very little variation from province to province.

* The expected percentage is 1% under the null hypothesis of an adequate fit. The residuals are differences between the sample covariance matrix and the covariance matrix fitted via the appropriate factor model. While all provinces have a higher than expected percentage of normalised residuals greater than three, man of theses are for variables for which there are large measurement errors.

The maximum value of the adjusted goodness of fit statistic is one, with high values indicating better fits.

TABLE A-n

Unconstrained Factor Analysis - Factor Loading Matrices
One factor for each province

	Ad	HB	Tara	Wgtn	Marl	Nelson	Cant	W'd	Otago
Violence	0.594	0.672	0.428	0.796	0.600	1.083	0.637	0.879	0.597
Drunkenn's	0.535	0.395	0.204	0.265	0.183	0.490	0.292	0.217	0.526
Overseas b.	-0.144	-0.202	-0.155	-0.149	-0.201	-0.234	-0.259	-0.285	-0.274
Irish born	0.637	0.474	0.352	0.287	0.274	0.592	0.437	0.858	0.369
Adult males	-0.031	0.041	-0.013	-0.039	0.054	0.134	0.009	0.153	0.048
Adult females	-0.165	-0.187	-0.142	-0.190	-0.170	-0.161	-0.190	-0.151	-0.211
Ratio men to women	0.115	0.245	0.124	0.131	0.291	0.397	0.168	0.398	0.251
Urban pop.	-0.141	-0.125	-0.007	-0.169	-0.067	0.035	-0.204	-0.137	-0.147
Males 21-40	0.058	0.195	0.043	0.046	0.204	0.326	0.170	0.173	0.202
Small dwllgs	0.324	0.449	0.326	0.429	0.485	0.847	0.392	1.283	0.619
Total dwllgs	-0.013	-0.033	-0.011	0.011	0.002	0.059	-0.039	0.178	-0.014
Ratio small to total dwllgs	0.343	0.517	0.350	0.392	0.498	0.678	0.462	0.802	0.649
Spirits cons'n	0.321	0.619	0.114	0.437	0.399	0.437	0.445	0.986	0.682
Beer cons'n	0.087	0.190	0.008	0.162	0.293	0.308	0.275	0.083	-0.126
£ imports	-0.522	-0.032	-0.194	-0.790	-0.021	0.300	-0.192	0.251	0.013
£ exports	-0.418	-0.629	-0.790	-0.685	-0.167	0.172	-0.278	0.096	-0.108
Manufact'g HP	-0.962	-0.675	-0.724	-0.921	-0.214	-0.290	-0.608	-1.281	-0.778
Manufact'g Employees	-0.253	-0.211	-0.193	-0.372	0.096	-0.124	-0.396	-0.600	-0.350
Police numbers	-0.015	0.188	0.199	0.086	0.170	0.286	0.212	0.216	0.169
Immigration	0.072	0.749	0.128	-0.462	0.102	0.325	0.981	0.642	0.923
Emigration	-0.120	0.086	0.028	-1.205	0.010	0.147	0.214	1.286	0.402
Civil suits	0.040	0.215	0.266	0.359	0.454	0.683	0.611	0.703	0.432

Notes: For description of variables see appendix 1. Except for ratio data, all variables are expressed as proportions of population, or as rates per 100,000.
Although factor loadings are tabulated relative to the variables for which they are loadings, the table refers to factor loadings rather than variables, per se.

TABLE A-m

Unconstrained Factor Analysis -
Correlation of variables with factor score for each province
One factor for each province

	Ad	HB	Tara	Wgtn	Marl	Nelson	Cant	W'd	Otago
Violence	0.872	0.848	0.816	0.927	0.723	0.959	0.932	0.792	0.931
Drunkenn's	0.791	0.685	0.616	0.545	0.505	0.855	0.477	0.480	0.947
Overseas b.	0.897	0.958	0.975	0.921	0.999	0.966	0.968	0.986	0.996
Irish born	0.988	0.987	0.996	0.839	0.902	0.976	0.884	0.997	0.940
Adult males	-0.510	0.556	-0.044	-0.428	0.485	0.855	0.215	0.873	0.568
Adult females	-0.980	-0.952	-0.924	-0.869	-0.909	-0.856	-0.855	-0.925	-0.943
Ratio men to women	0.953	0.975	0.954	0.935	0.962	0.993	0.999	0.975	0.980
Urban pop.	-0.752	-0.893	0.188	-0.899	-0.584	0.370	-0.912	-0.750	-0.848
Males 21-40	0.466	0.900	0.442	0.554	0.812	0.868	0.860	0.344	0.830
Small dwllgs	0.880	0.863	0.974	0.687	0.916	0.979	0.972	0.986	0.974
Total dwllgs	-0.311	-0.702	-0.031	0.201	-0.008	0.843	-0.379	0.984	-0.238
Ratio small to total dwllgs	0.914	0.860	0.949	0.827	0.955	0.948	0.970	0.985	0.992
Spirits cons'n	0.813	0.845	0.615	0.896	0.915	0.930	0.859	0.910	0.932
Beer cons'n	0.518	0.762	0.077	0.756	0.726	0.850	0.708	-0.078	-0.103
£ imports	-0.709	-0.188	-0.607	-0.453	-0.324	0.905	-0.277	0.550	0.104
£ exports	-0.854	-0.798	-0.720	-0.650	-0.459	0.656	-0.649	0.047	-0.202
Manufact'g HP	-0.840	-0.667	-0.754	-0.665	-0.618	-0.682	-0.634	-0.759	-0.754
Manufact'g Employees	-0.826	-0.825	-0.838	-0.945	0.424	-0.333	-0.910	-0.880	-0.942
Police numbers	-0.118	0.830	0.730	0.728	0.548	0.894	0.783	0.544	0.802
Immigration	0.113	0.483	0.184	-0.068	0.255	0.546	0.677	0.709	0.699
Emigration	-0.180	0.409	0.123	-0.802	0.079	0.721	0.664	0.788	0.380
Civil suits	0.274	0.578	0.527	0.700	0.832	0.936	0.882	0.737	0.869

Note: For description of variables see appendix 1. Except for ratio data, all variables are expressed as proportions of population or rates per 100,000.

Historical Social Research, Vol. 16 — 1991 — No. 3, 40-68

TABLE A-IV

Constrained Factor Analysis - Factor Loading Matrices
One factor for each province

	Ad	HB	Tara	Wgtn	Marl	Nelson	Cant	W'd	Otago
Violence	0.393	0.546	0.337	0.283	0.357	0.617	0.560	1.030	0.657
Drunkenness	0.159	0.223	0.137	0.115	0.146	0.251	0.228	0.420	0.268
Overseas b.	-0.128	-0.179	-0.110	-0.092	-0.117	-0.202	-0.183	-0.337	-0.215
Irish born	0.324	0.450	0.278	0.233	0.295	0.509	0.462	0.850	0.542
Adult males	0.050	0.070	0.043	0.036	0.045	0.079	0.071	0.131	0.084
Adult females	-0.083	-0.115	-0.071	-0.060	-0.075	-0.130	-0.118	-0.217	-0.139
Ratio men to women	0.160	0.223	0.138	0.115	0.146	0.252	0.229	0.422	0.269
Urban pop.	-0.059	-0.083	-0.051	-0.043	-0.054	-0.094	-0.085	-0.156	-0.099
Males 21-40	0.093	0.120	0.080	0.067	0.085	0.146	0.133	0.244	0.155
Small dwllgs	0.452	0.632	0.390	0.327	0.413	0.714	0.648	1.192	0.760
Total dwllgs	0.043	0.060	0.037	0.031	0.039	0.067	0.061	0.113	0.072
Ratio small total dwllgs	0.338	0.470	0.290	0.243	0.307	0.531	0.482	0.886	0.565
Spirits cons'n	0.369	0.514	0.317	0.266	0.336	0.580	0.527	0.969	0.618
Beer cons'n	0.053	0.074	0.046	0.038	0.048	0.084	0.076	0.139	0.089
£ imports	0.053	0.074	0.046	0.038	0.048	0.084	0.076	0.139	0.089
£ exports	-0.030	-0.042	-0.026	-0.022	-0.028	-0.048	-0.043	-0.080	-0.051
Manufact'ing HP	-0.455	-0.633	-0.391	-0.327	-0.414	-0.716	-0.650	-1.195	-0.762
Manufact'ing Employees	-0.210	-0.293	-0.180	-0.152	-0.191	-0.331	-0.301	-0.553	-0.352
Police numbers	0.101	0.140	0.087	0.073	0.092	0.160	0.144	0.266	0.170
Immigration	0.310	0.431	0.266	0.223	0.281	0.487	0.442	0.813	0.519
Emigration	0.345	0.479	0.296	0.248	0.316	0.542	0.492	0.904	0.578
Civil suits	0.285	0.396	0.244	0.205	0.259	0.447	0.406	0.747	0.476

Notes: For description of variables see appendix 1. Except for ratio data, all variables are expressed as proportions of population, or as rates per 100,000.
Although the factor loadings are tabulated relative to the variables for which they are loadings, the table refers to factor loadings rather than variables, per se.

TABLE A-V

Constrained Factor Analysis -
Correlation of variables with factor score for each province
One factor for each province

	Ad	HB	Tara	Wgtn	Marl	Nelson	Cant	W'd	Otago
Violence	0.910	0.866	0.793	0.965	0.730	0.952	0.944	0.853	0.952
Drunkenn's	0.897	0.685	0.668	0.627	0.526	0.840	0.505	0.545	0.944
Overseas b.	0.954	0.951	0.959	0.911	0.982	0.969	0.990	0.989	0.985
Irish born	0.960	0.977	0.984	0.865	0.852	0.979	0.911	0.981	0.902
Adult males	-0.295	0.584	-0.037	-0.337	0.624	0.847	0.147	0.903	0.646
Adult females	-0.891	-0.928	-0.894	-0.855	-0.869	-0.863	-0.880	-0.891	-0.907
Ratio men to women	0.979	0.971	0.965	0.951	0.989	0.988	0.992	0.981	0.990
Urban pop.	-0.617	-0.877	0.217	-0.851	-0.699	0.344	-0.943	-0.767	-0.867
Males 21-40	0.630	0.905	0.513	0.657	0.902	0.863	0.838	0.463	0.879
Small dwllgs	0.951	0.881	0.963	0.865	0.957	0.986	0.972	0.997	0.987
Total dwllgs	-0.103	-0.661	0.053	0.416	0.144	0.845	-0.409	0.979	0.141
Ratio small to total dwllgs	0.968	0.877	0.926	0.953	0.971	0.958	0.969	0.984	0.995
Spirits cons'n	0.895	0.859	0.688	0.870	0.944	0.934	0.870	0.953	0.953
Beer cons'n	0.540	0.730	0.072	0.680	0.713	0.844	0.674	0.042	0.029
£ imports	-0.557	-0.166	-0.568	-0.490	-0.254	0.897	-0.327	0.635	0.171
£ exports	-0.716	-0.772	-0.705	-0.684	-0.409	0.640	-0.687	0.111	-0.132
Manufact'g HP	-0.716	-0.650	-0.739	-0.662	-0.537	-0.714	-0.691	-0.719	-0.710
Manufact'g Employees	-0.908	-0.861	-0.853	-0.922	0.393	-0.350	-0.940	-0.838	-0.950
Police numbers	-0.172	0.822	0.706	0.729	0.607	0.895	0.793	0.629	0.855
Immigration	0.090	0.460	0.209	-0.173	0.276	0.556	0.678	0.760	0.674
Emigration	-0.243	0.396	0.101	-0.717	0.071	0.726	0.687	0.852	0.333
Civil suits	0.406	0.611	0.585	0.773	0.858	0.923	0.889	0.816	0.885

Notes: For description of variables see appendix. Except for ratio data, all variables are expressed as proportions of population, or as rates per 100,000.

Historical Social Research, Vol. 16 — 1991 — No. 3, 40-68

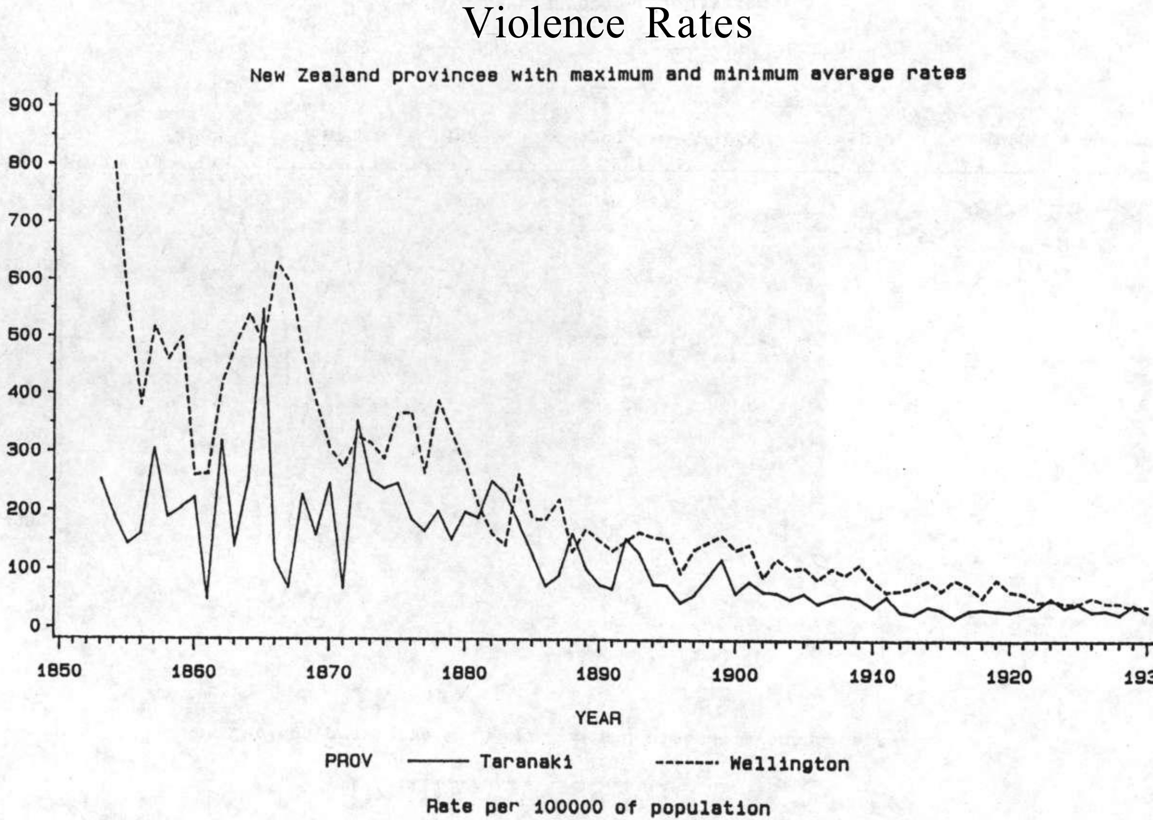
TABLE A-VI

Standard errors for difference between factor loading matrices
- Unconstrained less constrained
One factor per province

	Ad	HB	Tara	Wgtn	Marl	Nelson	Cant	W'd	Otago
Violence	0.074	0.085	0.068	0.090	0.097	0.112	0.064	0.114	0.070
Drunkenn's	0.071	0.064	0.043	0.086	0.042	0.063	0.058	0.051	0.055
Overseas b.	0.017	0.021	0.017	0.015	0.020	0.023	0.024	0.027	0.026
Irish born	0.063	0.049	0.038	0.031	0.037	0.056	0.049	0.081	0.040
Adult males	0.009	0.010	0.009	0.008	0.012	0.017	0.010	0.017	0.011
Adult females	0.017	0.019	0.015	0.019	0.020	0.019	0.021	0.018	0.022
Ratio men to women	0.014	0.024	0.015	0.015	0.029	0.038	0.017	0.039	0.025
Urban pop.	0.020	0.014	0.015	0.020	0.013	0.013	0.020	0.019	0.017
Males 21-40	0.016	0.023	0.017	0.014	0.026	0.041	0.023	0.044	0.026
Small dwllgs	0.040	0.057	0.038	0.075	0.057	0.081	0.043	0.118	0.063
Total dwllgs	0.006	0.005	0.007	0.015	0.009	0.008	0.010	0.017	0.007
Ratio small to total dwllgs	0.038	0.064	0.039	0.050	0.054	0.067	0.047	0.076	0.063
Spirits cons'n	0.044	0.076	0.035	0.062	0.045	0.048	0.054	0.101	0.076
Beer cons'n	0.024	0.030	0.020	0.033	0.046	0.041	0.048	0.102	0.084
£ imports	0.084	0.038	0.038	0.137	0.025	0.040	0.086	0.051	0.074
£ exports	0.055	0.086	0.120	0.092	0.054	0.039	0.049	0.072	0.048
Manufact'g HP	0.133	0.120	0.109	0.132	0.063	0.062	0.108	0.241	0.115
Manufact'g Employees	0.034	0.030	0.031	0.039	0.030	0.047	0.041	0.074	0.039
Police numbers	0.019	0.023	0.032	0.018	0.038	0.033	0.028	0.043	0.025
Immigration	0.098	0.215	0.097	0.203	0.058	0.082	0.169	0.101	0.149
Emigration	0.122	0.042	0.042	0.166	0.041	0.039	0.044	0.167	0.116
Civil suits	0.039	0.052	0.095	0.055	0.065	0.075	0.078	0.097	0.059

Notes: For description of variables see appendix 1. Except for ratio data, all variables are expressed as proportions of population, or as rates per 100,000.
Although the particular factor loading standard errors are tabulated relative to the variables for which they are loadings, the table refers to factor loadings rather than variables, per se.

Historical Social Research, Vol. 16 — 1991 — No. 3, 40-68

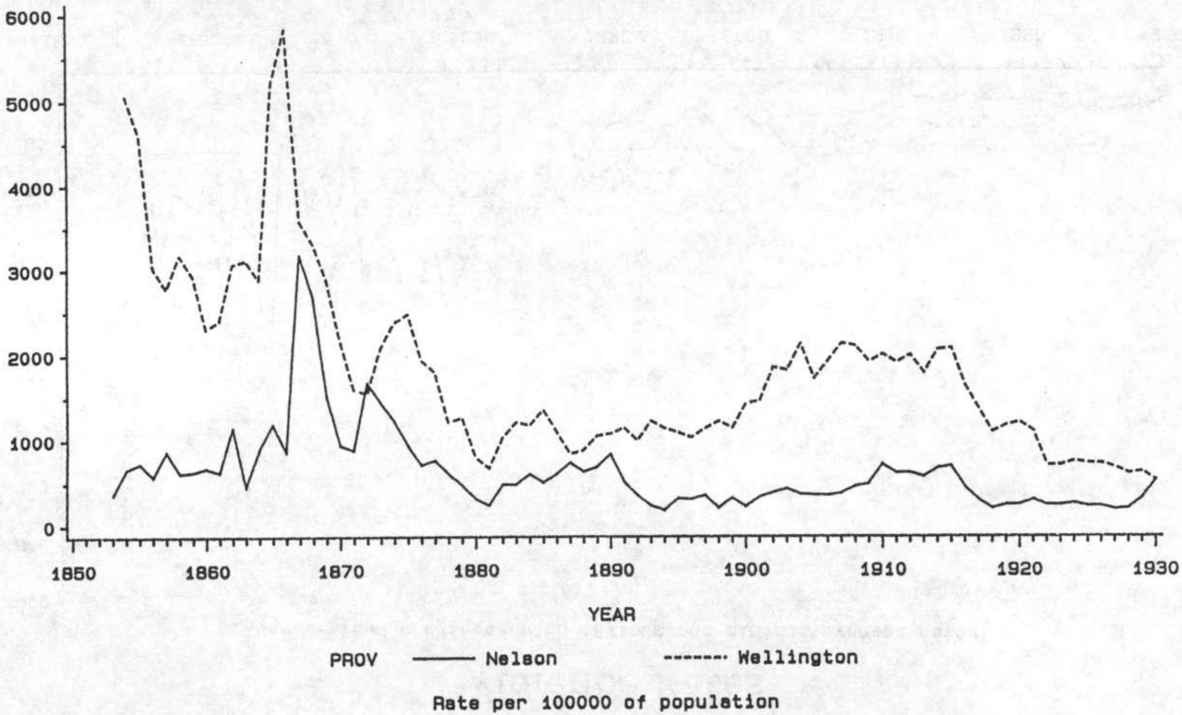


Historical Social Research, Vol. 16 — 1991 — No. 3, 40-68

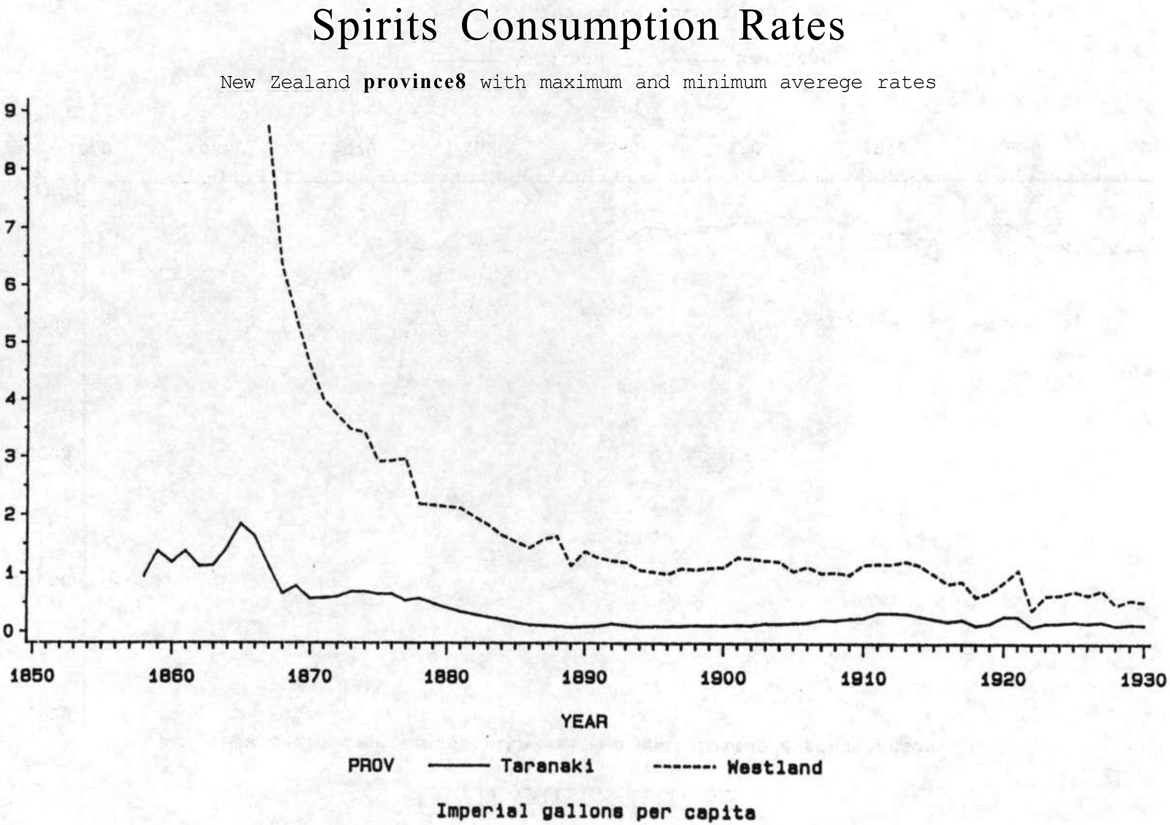
No 2

Drunkenness Rates

New Zealand provincee with maximum and minimum average ratee



Historical Social Research, Vol. 16 — 1991 — No. 3, 40-68

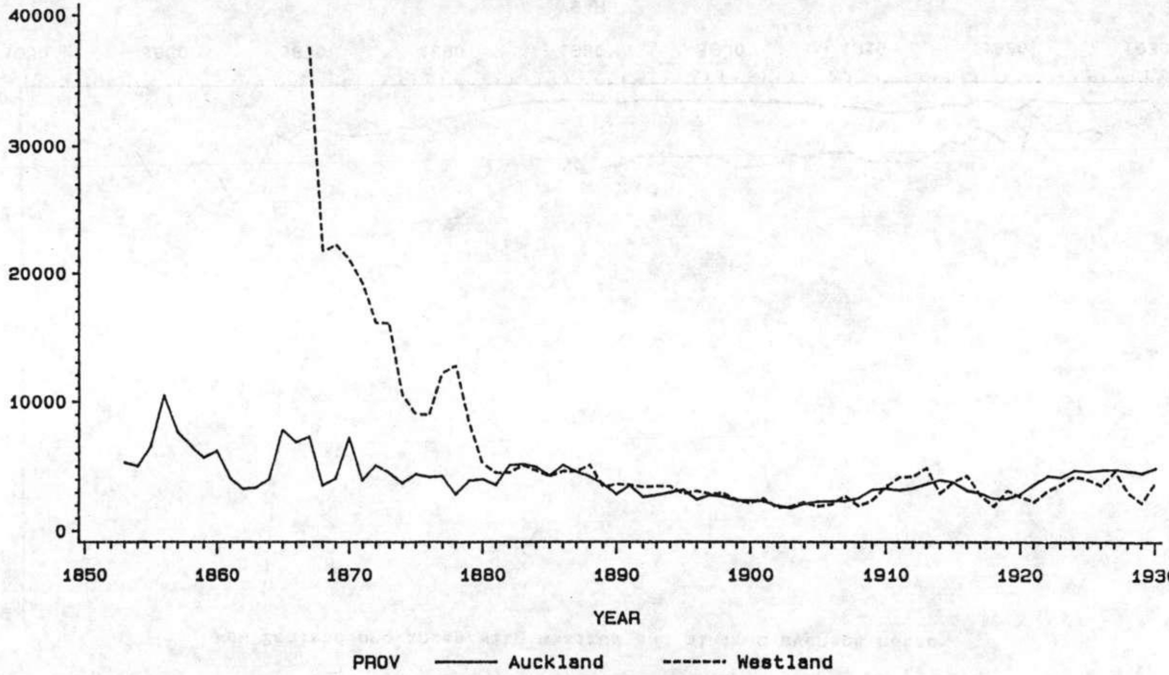


as

Historical Social Research, Vol. 16 — 1991 — No. 3, 40-68

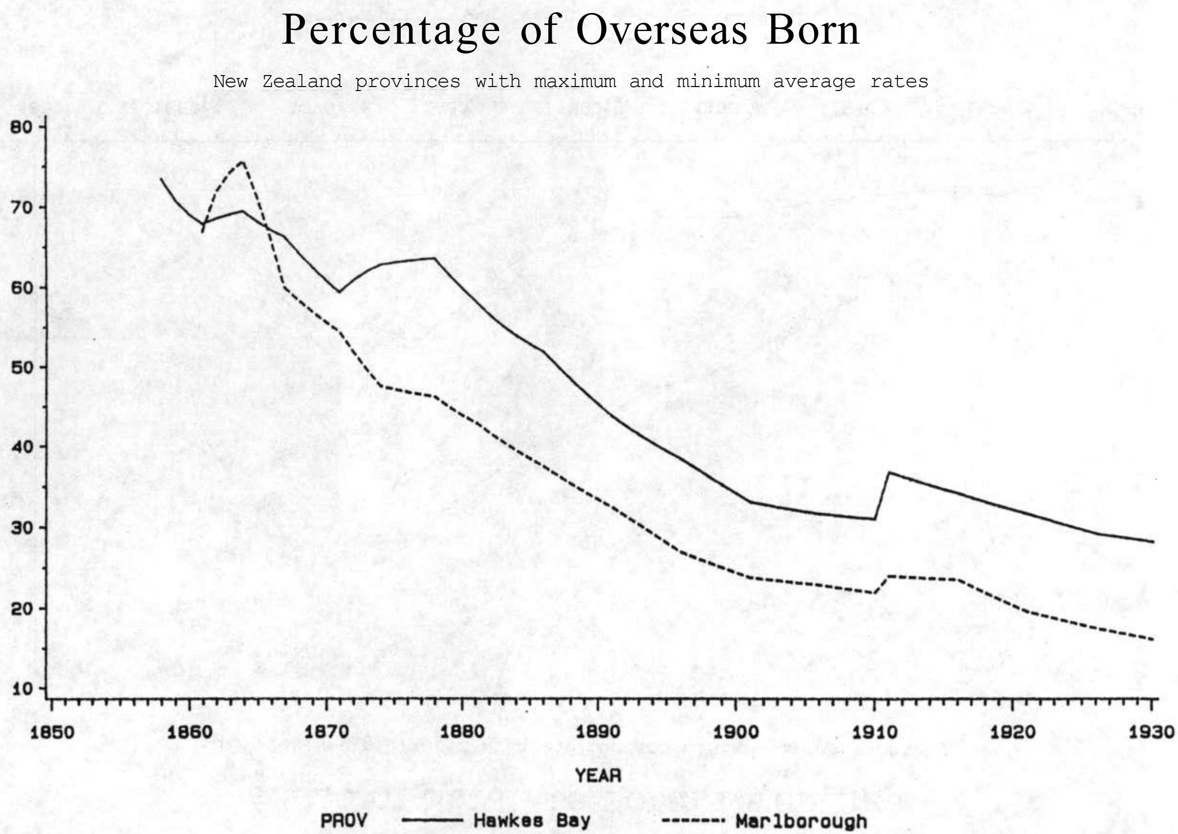
Civil Suits Rates

New Zealend provincee with maximum and minimum average ratee



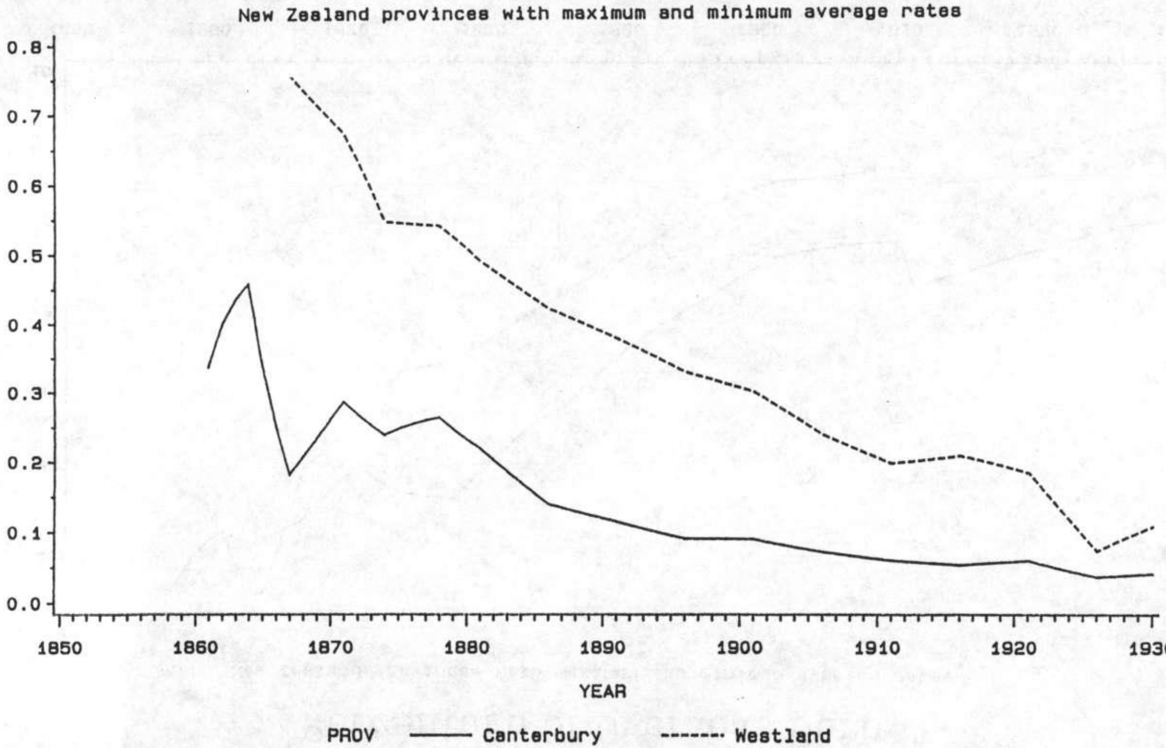
Rate prr 100000 of population

Historical Social Research, Vol. 16 — 1991 — No. 3, 40-68



Historical Social Research, Vol. 16 — 1991 — No. 3, 40-68

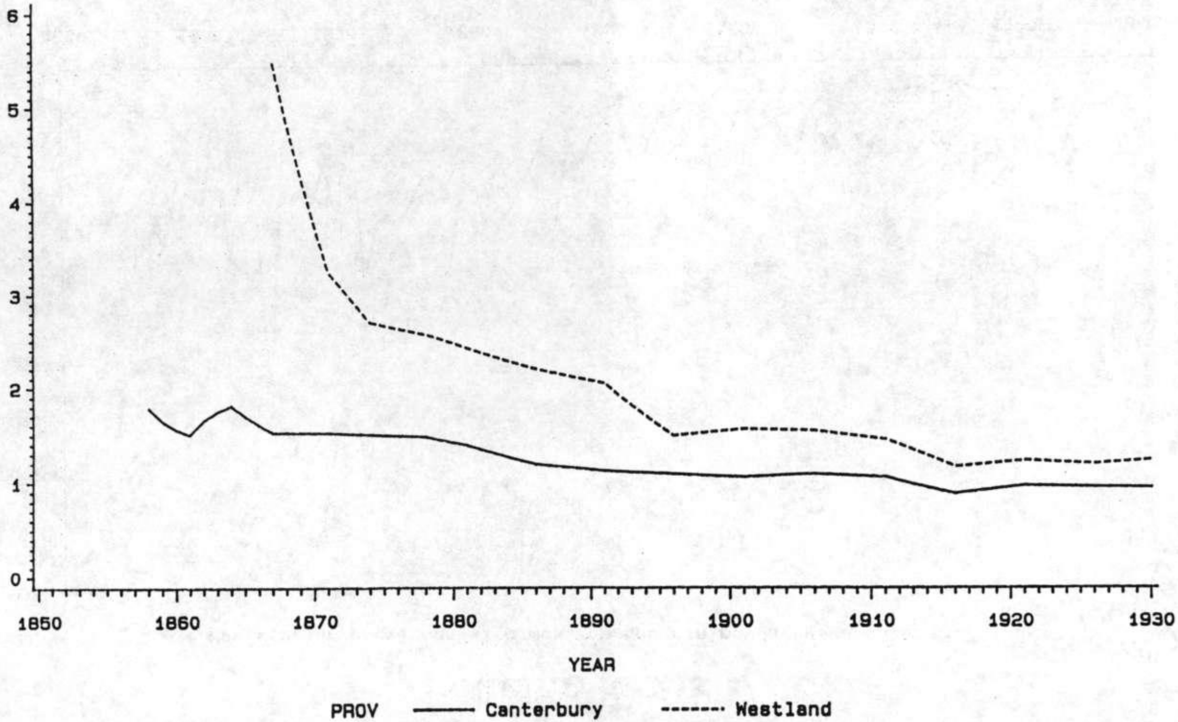
Ratio of Small to Total Dwellings



3

Ratio of Adult Males to Adult Females

New Zealand provinces with maximum and minimum average ratea



Historical Social Research, Vol. 16 — 1991 — No. 3, 40-68

Policing Rates